A TOOL FOR DETERMINING IMAGE DISCRIMINABILITY

Michael S. Landy[†]
Department of Psychology & Center for Neural Science
New York University

We describe *discrim*, a Matlab-based program for investigating image discriminability in various display systems. *Discrim* allows the user to manipulate images representing scene content. These images can be built up out of image primitives (sine wave gratings, Gabor patches) and image files from other sources. The user may then manipulate a model of the sensor and display characteristics. The sensor and display model consists of four stages: an initial spatial filter preceding the sensor, Poisson noise, a point nonlinearity ("Gamma") and output noise. The user may view the effects each of these stages has on the image. Finally, the user may specify a pair of images as input to an image discriminability model.

INTRODUCTION

In recent years, a number of computational models have been proposed for early visual coding, leading to models of various visual tasks such as detection, discrimination and classification. In this paper, we describe a program, *discrim*, that was originally written to allow the user to manipulate images and apply image discrimination models to them. It has been substantially revised recently, and now includes the ability to model both the human observer as well as the sensor and display system that is used to detect and display the image materials (e.g., night vision equipment). *Discrim* is fully described in a user's manual. Here, we concentrate on the capabilities of the program as well as some general design issues.

Display system model
Image sensors and display systems have a number of characteristics that can impact the ability of human observers to use the resulting displayed images. This is particularly true in challenging sensor environments such as the low-light conditions that require the use of night-vision equipment. In such conditions, visual displays are photon-limited. To contend with the lack of light, sensor systems must lower spatial resolution (effectively integrating over larger areas of the image), lower temporal resolution (integrating over time) and/or become more quantum efficient. All of these factors have an effect on image quality and hence on

human performance using the resulting displays. Image sensors and display systems are typically modeled as a series of image manipulations including spatial filters, point nonlinearities and corruption by noise. *Discrim* implements a simple sensor and display model consisting of four stages. These are, in order of application, (1) a linear spatial filter, (2) Poisson input noise, (3) a point nonlinearity, and (4) Gaussian output noise.

Spatial filter
The input spatial filter may be used to mimic the effects of the optics of the image sensor. *Discrim* can apply several types of filters. One type is a 2-dimensional, Cartesian-separable Gaussian. Very loosely speaking, Cartesian-separable means that the filter may be specified as a product of a filter applied to the horizontal frequencies multiplied by one applied to the vertical frequencies. A second type is a Cartesian-separable filter with the vertical and horizontal modulation transfer functions supplied in a file. Finally, the user can specify a filter that is difference of Gaussians, each of which is Cartesian-separable.

Poisson input noise
The user may specify, or have the program compute, the mean quantum catch of the individual sensor pixels. When that quantum catch is low, as it must be in the low-light conditions for which night vision equipment is designed, the effects of the Poisson statistics of light become important. As an example, Fig. 1

shows a 128×128 image of a sine wave grating with Poisson noise, assuming an average quantum catch of 10 photons per pixel.

Point nonlinearity
The user may specify the nonlinearity applied to individual pixels. Typically, both image sensors (film, vidicons, etc.) and image displays (CRTs, in particular) are characterized by a so-called "gamma curve". *Discrim* includes a single nonlinearity in its sensor and display model. One can think of it as a lumped model that combines both the sensor and display nonlinearities. We implement a minor generalization of the gamma curve by allowing for a level below which no output occurs (which we call "liftoff"). Thus, the output of the nonlinearity is $y = \lfloor x - x_0 \rfloor^\gamma$, where x is the input level, x_0 is the liftoff level, and γ is the degree of nonlinearity (2.3-3 is a typical range of values for a CRT system, and such values are often built into devices, such as DLP projectors, which don't have an implicit pixel nonlinearity). Fig. 2 shows an example of the window used in *discrim* to specify the current gamma function.
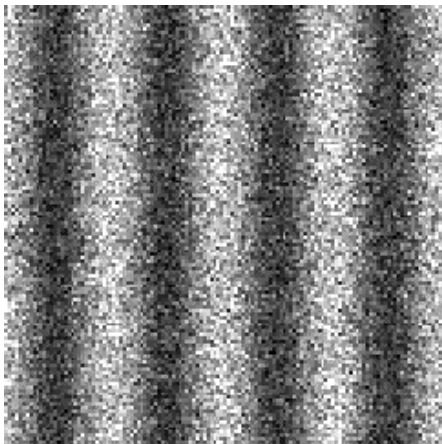

Figure 1. A sine wave grating corrupted by Poisson noise.

Output noise
    Finally, *discrim* includes the option of having Gaussian noise added to the displayed pixels after the point nonlinearity has been applied. This may be used to model imperfections in the display device, but can be
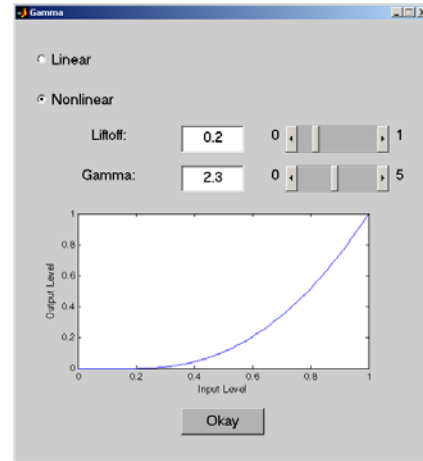

Figure 2. Specification of the gamma curve in *discrim.*

used for other modeling applications as well (e.g., models of medical imaging devices).

Observer model
In recent years, a number of computational models have been proposed for early visual coding. These models are based on known facts of the early architecture of the human visual system as well as taking into account empirical observations of human performance in visual discrimination tasks.

Several aspects of the visual system are typically reflected in vision models. They may include any or all of the following. (1) *Spatial channels.* Early in the visual system, patterns are coded using channels tuned for spatial frequency and orientation. Typical bandwidth estimates are one octave in spatial frequency and 30º in orientation. Peak sensitivity varies with spatial frequency, and the envelope of channel sensitivities results in the human contrast sensitivity function. (2) *Retinal inhomogeneity and sampling.* The peak sensitivity, range of spatial frequencies coded and density of sampling varies as one proceeds from the fovea to peripheral locations. (3) *Masking.* Target patterns are generally, although not always, more visible on a uniform background than in the presence of other patterned visual input. By and large, current visual models result in a vector of responses from the modeled channels that serves as the code for the input image. Vision models may be used to predict the results of typical tasks faced by the visual system,

2

including (1) *detection*: visibility of a pattern against a uniform background; (2) *discrimination*: ability to detect the difference between two images, which includes detection of a target against a non-uniform background; and (3) *classification*: determining whether a pattern, viewed alone or on a noisy background, is one of *n* possible test patterns (e.g., letter identification). At present, *discrim* includes a single image discrimination model called the "single filter, uniform masking" (SFUM) model (Ahumada, 1996; Ahumada & Beard, 1996, 1997a,b; Rohaly, Ahumada & Watson, 1997). The model includes a contrast sensitivity function as well as a simple model of pattern masking. SFUM was designed to estimate the value of d' for discriminating two given, fixed images. For example, in evaluating a lossy image compression scheme, SFUM will provide an estimate of the ability of an observer to discriminate an original image from its compressed, distorted counterpart. The resulting d' (pronounced "d prime") value indicates the degree of discriminability. A d' value of zero indicates the two images are completely indiscriminable, so that an observer would be 50% correct (i.e., guessing) on a two-alternative forced-choice (2AFC) task. d' values of 1 and 2 correspond to performance of 76% and 92% correct in a 2AFC task, respectively.

The *SFUM* model was designed to estimate the discriminability of a pair of fixed images. The display model, described in the previous section, involves two possible sources of noise: input Poisson noise and output Gaussian noise. When either or both of those noise sources are enabled, the intent of the *discrim* program is to allow the user to estimate the discriminability of the two input images under conditions of stochastic variability due to the noise source(s) (and other image distortions). That is, on any given trial an observer will see a different retinal image due to the variability of the noise from trial to trial. If we simply added different, independent Poisson and/or Gaussian noise samples to each of the two images and then applied the SFUM model, SFUM would attempt to estimate discriminability not only of the underlying images, but of the two noise samples as well. Clearly, this is not appropriate. What is of interest is the observer's ability to discriminate the underlying scenes despite the noise, not their ability to discriminate the noise samples.

Thus, the SFUM model is not well-suited to the problem at hand. However, we have implemented the SFUM model in a way that should allow it to provide reasonable estimates. We do this by using the same sample of noise, Poisson and/or Gaussian, for both images. Thus, there are no differences between the two noise samples that artificially inflate the d' estimates. Gaussian noise is an additive process that is independent of the image content. It is a simple matter to generate a Gaussian noise image, and add it to both input images. On the other hand, Poisson noise depends on the image content. The variance of the noise added to any given pixel is equal to the value of that pixel. This means that the use of the same noise image for both input images is not an accurate reflection of Poisson statistics. We have settled on an approximation that we feel is adequate for the sorts of threshold detection tasks for which *discrim* is most appropriate. When Poisson noise is used with the SFUM model, the two input images are first blurred using the current spatial filter. The blurred images are averaged, and that average image is subjected to Poisson noise. The difference between the noisy image and the average image (the *error image*) is treated as an additive noise source. That error image is then added to the individual blurred images to simulate a Poisson noise source that perturbs both images identically. Note that each time *discrim* calculates a d' value, new samples of Poisson and/or Gaussian noise are used. Thus, the user can average over several such calculations to guard against an outlier value due to an atypical noise sample.

The *discrim* program is designed to be able to include additional vision models. In particular, there is a large literature (mostly from the medical imaging community) of visual detection and discrimination models for visual targets in patterned and noisy backgrounds (Barrett, Yao, Rolland & Myers, 1993; Bochud, Abbey & Eckstein, 2000; Burgess, 1999; Burgess, Li & Abbey, 1997; Eckstein et al., 2003; King, de Vries & Soares, 1997; Myers et al., 1985; Rolland & Barrett, 1992; Wagner & Weaver,

1972; for reviews see Eckstein, Abbey & Bochud, 2000; Wagner & Brown, 1985). These models provide an estimate of d' given the input images and descriptions of the noise (variance, spatial correlation, etc.). Thus, for these models, *discrim* is already set up to provide the required information, and the issue of using identical noise samples for the two input images shouldn't arise.

Discrim: a tool for modeling image discriminability

Fig. 3 shows the main window of *discrim*. The program is written in the Matlab language. It provides a number of capabilities in the main window, through menus, and through pop-up subwindows that are invoked via a menu. The user builds up a library of input images that are either read in from files or constructed using various built-in image manipulation primitives. On the main window, two of these images can be displayed (here, two slightly different airport images are shown). The currently active image is distinguished by its magenta outline. A button below the images allows the user to request that *discrim* calculate d' for discriminating the two images currently displayed, using the current parameters that govern the display and observer models.

There are six menus available. The *Model* menu is for choosing the image discriminability model to be simulated. Currently, SFUM is the only available choice.

The *Model Parameters* menu brings up a pop-up window for setting the parameters for each available image discrimination model. Currently, the only choice is to invoke a window for setting the parameters of the SFUM model. These include parameters controlling its contrast sensitivity function and the degree of masking.

The *Image* menu allows the user to create, delete, load and save images. Note, in *discrim* the library of images that are created and manipulated are *input* images. That is, they are the raw, undistorted images describing the scene presented to the display model. These images are undistorted by the display filter, Poisson input noise, gamma curve or output noise.
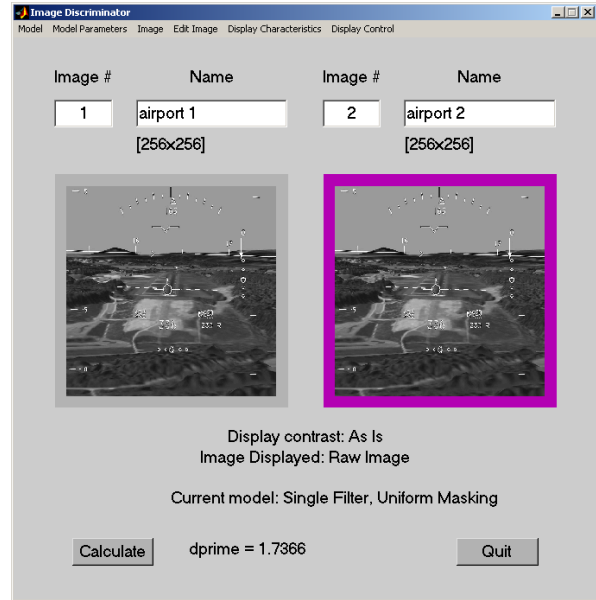

Figure 3. The main window of *discrim*.

The *Edit Image* menu allows the user to manipulate the images. One can add a pattern to an image (sine wave gratings, Gabor patches). Two images can be combined (e.g., adding a target to a background). An image can have its contrast scaled and have extreme pixel values clipped.

The *Display Characteristics* menu is used to control the parameters of the image display model. It is used to invoke pop-up windows for each stage in the display model, including the initial spatial filter, the input Poisson noise, the gamma nonlinearity and the output noise. An additional window allows the user to specify the viewing geometry, including the size of the image in pixels and its visual size.

In Fig. 3, the images shown are input images, undistorted by the various stages of the display model. In the *Display Control* menu, the user can specify which form of image distortion they would like to view.

4

This includes the undistorted image, as shown, as well as the image after each display distortion has been effected, in turn.

CONCLUSIONS

Discrim is a tool that is easy to use. It is our hope that display modelers and evaluators will be able to use *discrim* to test the quality of displays and their usefulness for particular visual tasks. We are making the code and documentation freely available to the general public at http://www.cns.nyu.edu/~msl/discrim. We hope that others will make use of the software, and will let us know what other capabilities would be useful. Clearly, the most important improvement would be to include additional models of image discrimination, especially those that make allowance for discrimination of noisy images.

REFERENCES

Ahumada, A. J., Jr. (1996). Simplified vision models for image quality assessment. In J. Morreale (Ed.), *SID International Symposium Digest of Technical Papers*, *27*, 397-400. Santa Ana, CA: Society for Information Display.

Ahumada, A. J., Jr. & Beard, B. L. (1996). Object detection in a noisy scene. In B. E. Rogowitz & J. Allebach (Eds.), *Human Vision, Visual Processing, and Digital Display VII*, *2657*, 190-199. Bellingham, WA: SPIE.

Ahumada, A. J., Jr. & Beard, B. L. (1997a). Image discrimination models predict detection in fixed but not random noise. *Journal of the Optical Society of America A*, *14*, 2471-2476.

Ahumada, A. J., Jr. & Beard, B. L. (1997b). Image discrimination models: Detection in fixed and random noise. In B. E. Rogowitz & T. N. Pappas (Eds.), *Human Vision, Visual Processing, and Digital Display VIII*, *3016*, 34-43. Bellingham, WA: SPIE.

Barrett, H. H., Yao, J., Rolland, J.P. & Myers, K. J. (1993). Model observers for assessment of image quality. *Proceedings of the National Academy of Sciences USA*, *90*, 9758- 9765.

Bochud, F. O., Abbey, C. A. & Eckstein, M. P. (2000). Visual signal detection in structured backgrounds III, Calculation of figures of merit for model observers in non-stationary backgrounds. *Journal of the Optical Society of America A*, *17*, 193-205.

Burgess, A. E. (1999). Visual signal detection with two-component noise: low-pass spectrum effect. *Journal of the Optical Society of America A*, *16*, 694-704.

Burgess, A. E., Li, X. & Abbey, C. K. (1997). Visual signal detectability with two noise components: anomalous masking effects. *Journal of the Optical Society of America A*, *14*, 2420-2442.

Eckstein, M. P., Abbey, C. K. & Bochud, F. O. (2000). A practical guide to model observers for visual detection in synthetic and natural noisy images. In J. Beutel, H. L. Kundel & R. L. van Metter (Eds.), *Handbook of Medical Imaging, Vol. 1, Physics and Psychophysics* (pp. 593-628). Bellingham, WA: SPIE Press.

Eckstein, M. P., Bartroff, J. L., Abbey, C. K., Whiting, J. S., Bochud, F. O. (2003). Automated computer evaluation and optimization of image compression of x-ray coronary angiograms for signal known exactly tasks. *Optics Express*, *11*, 460-475. http://www.opticsexpress.org/abstract.cfm?URI=OPEX-11-5-460

King, M. A., de Vries, D. J. & Soares, E. J. (1997). Comparison of the channelized Hotelling and human observers for lesion detection in hepatic SPECT imaging. *Proc. SPIE Image Perc.*, *3036*, 14-20.

Myers, K. J., Barrett, H. H., Borgstrom, M. C., Patton, D. D. & Seeley, G. W. (1985). Effect of noise correlation on detectability of disk signals in medical imaging. *Journal of the Optical Society of America A*, *2*, 1752-1759.

Rohaly, A. M., Ahumada, A. J., Jr. & Watson, A. B. (1997). Object detection in natural backgrounds predicted by discrimination performance and models. *Vision Research*, *37*, 3225-3235.

Rolland, J. P. & Barrett, H. H. (1992). Effect of random inhomogeneity on observer detection performance. *Journal of the Optical Society of America A*, *9*, 649-658.

Wagner, R. F. & Brown, D. G. (1985). Unified SNR analysis of medical imaging systems. *Phys. Med. Biol.*, *30*, 489-518.

Wagner, R. F. & Weaver, K. E. (1972). An assortment of image quality indices for radiographic film-screen combinations – can they be resolved? *Appl. of Opt. Instr. in Medicine*, *Proc. SPIE*, *35*, 83-94.

[†] Address correspondence to: Michael S. Landy, Department of Psychology, New York University, 6 Washington Place, Rm. 961, New York, NY 10003, landy@nyu.edu, (212) 998-7857, fax: (212) 995-4349.